Front. Electr. Electron. Eng. China 2011, 6(2): 291–299
DOI 10.1007/s11460-011-0143-1

RESEARCH ARTICLE

Chaoxu MU, Changyin SUN

Data-based intelligent modeling and control for nonlinear systems

Abstract With the ever increasing complexity of industrial systems, model-based control has encountered difficulties and is facing problems, while the interest in data-based control has been booming. This paper gives an overview of data-based control, which divides it into two subfields, intelligent modeling and direct controller design. In the two subfields, some important methods concerning data-based control are intensively investigated. Within the framework of data-based modeling, main modeling technologies and control strategies are discussed, and then fundamental concepts and various algorithms are presented for the design of a data-based controller. Finally, some remaining challenges are suggested.

Keywords offline and online data, intelligent modeling, data-based control, perspective

1 Introduction

Since the latter half of the 20th century, classical and modern control theory have developed in a number of areas and branches, such as system identification, adaptive control, robust control, optimal control, variable structure control, and stochastic control, and have achieved remarkable results in industrial processes, aerospace, military, and many other fields. Classical and modern control can be called as traditional control due to a common property that control is based on mathematic models. It follows two steps to a practical application, first describing dynamic systems and establishing corresponding models, and second designing controllers based on special models to stabilize processes, which can achieve good effect for simple systems.

Received July 15, 2010; accepted February 22, 2011

Chaoxu MU, Changyin SUN (✉)
School of Automation, Southeast University, Nanjing 210096, China
E-mail: cysun@seu.edu.cn

However, in recent years, industrial processes have undergone significant changes along with the rapid development of information technology [1,2]. Typical examples are the processes seen in steel making, material processing, chemical plants, and mineral processes, etc. The increasing scale, complex nonlinearities, time-delays and multiple variables, all of these make controlled systems increasingly sophisticated and also have brought about tremendous challenges to traditional control [3]. When traditional control, i.e., model-based control, comes across a sophisticated system, it is often hard to establish a mathematic model and guarantee the control reliability. From the perspective of application, a lot of practical issues, such as chemical processes, controllers are required to be low cost with optimal targets, while model-based control needs high cost to realize effective control by much exploration about the model and controller. Batch process is another example [4], different batches have different products, so it is unthinkable to model each batch and each product to improve quality and output [5]. Therefore, a global mathematic model is unlikely to be acquired, and a local mathematic model is also not accurate with increasing complex factors of industrial systems. Traditional control is becoming more and more powerless and has encountered problems and challenges.

A new phenomenon for industrial processes is that industrial processes generate and store a great deal of data. The explosive growth of data is even from terabytes to petabytes. Data means information. How to utilize offline and online data to achieve effective control has become a pressing issue and the concept of data-based control is presented resulting from the above problem. In a small number of existing literatures, it is also called data-driven control, modeless control, model-free control, and so on.

Data-based control is only in its infancy, but has attracted a lot of attention from foreign and domestic researchers. The Institute of Mathematics and its Application of Minnesota University convened a special

3-day seminar in 2002, named "IMA hot topics workshop: data-driven control and optimization". The National Natural Science Foundation of China held a special symposium on "data-based control, decision-making, scheduling and fault diagnosis" in 2008. Control theory and methods have significant developments along with the progress of science, from the classical control based on transfer function, to the modern control based on state-space model, as well as to intelligent control without the dependence on strict models. Intelligent control is an advanced stage of control theory and data-based control is an important branch of intelligent control. Figure 1 shows the various stages of control theory and main analysis and synthesis methods.

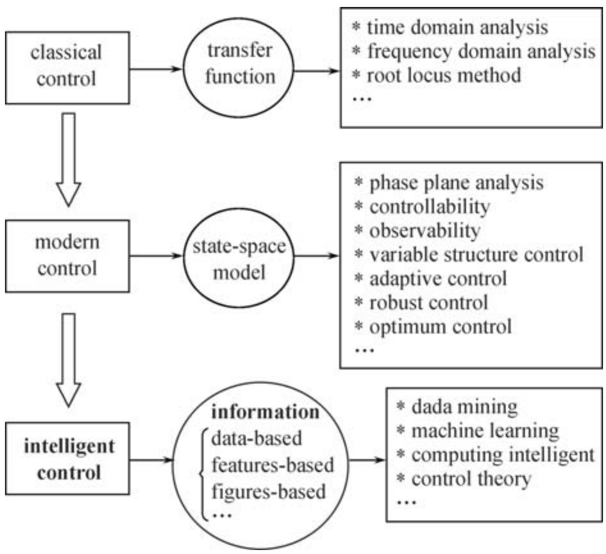


Fig. 1 Various stages of control theory and main methods

Data-based control has two major ideas. One is intelligent modeling based on data, and designing a corresponding controller according to the data-based model, then there are some researches about convergence, stability, and robustness under certain assumptions. The other idea is to design the controller directly by online and offline data and analyze properties of convergence, stability, and robustness. The main aim of this paper is to focus on the new data-based control by reviewing the main approaches adopted.

The paper is organized as follows. In Sect. 2, we discuss the present status and key problems of data-based intelligent modeling and control. Section 3 is devoted to direct methods from data to controllers. Finally, we conclude the paper and also present several considerations centered on data-based control in Sect. 4.

2 Data-based intelligent modeling and control

The performance of a controller depends on how well the system dynamics have been captured by the model. Hence, modeling is considered as a crucial technology so that accurate information about the system can be obtained and utilized.

2.1 Intelligent modeling methods based on data

A number of sensors in an industrial process record and store large volumes of data when the system works. Data-based modeling mainly explores the system modes from data beginning with data sources, to data preprocessing and data understanding, and last to modeling algorithms by technologies of data mining and machine learning.

Data mining and machine learning extract non-trivial, implicit, previously unknown, and potentially useful patterns or knowledge by automatically analyzing massive data. Classification, as a kind of form of data learning, has been widely studied and effective methods include decision trees induction, Bayesian network, support vector machines (SVMs), associative classification, neural networks (NNs), genetic algorithm, rough set, and so on [6–8]. Regression has its root in classification, so some revised versions of classification algorithms also have good performance in regression. Taking neural networks as an example, although the understanding of natural intelligent behavior and designing intelligent machines still remain a challenging topic [9,10], it has been widely used for modeling and control including back propagation (BP) networks, radial basis function (RBF) networks, Hopfield networks, self-organizing mapping (SOM) networks, counter propagation networks (CPNs), qubit neural networks, and so on [11]. SVM is another important methodology in the area of intelligent modeling, still also in classification. By solving convex optimization problems, it can find a global optimization result to overcome the local minimum in neural networks. It mainly includes ξ-support vector regression (SVR), v-SVR, least square-SVR, twin-SVR and ridge regression [12].

Load forecasting, as a typical modeling problem, is focused on economic generation of power, economic allocation between plants, maintenance scheduling and system security. We proposed a fuzzy two-stage modeling technique with particle swarm optimization based on data for short-term load forecasting in Ref. [13]. In order to build a more excellent model to forecast loads, we gave a novel hybrid algorithm combining the fuzzy SVR method with the linear extrapolation based on similar days in Ref. [14]. The fuzzy SVR was used to consider the lower load demands and the normal load in weekdays was forecasted by the linear extrapolation based on similar days.

In general, most regression algorithms are considered as supervised learning, which extracts information from existing label data and produces the correct output when a new input comes. Semi-supervised learning, as a rising modeling method, seeks to understand how to combine labeled and unlabeled data and take advantage of such a combination. In Refs. [15,16], semi-supervised SVM and semi-supervised least square (LS)-SVM were proposed. In Ref. [17], the authors presented a simple unification of several supervised and unsupervised training principles through the concept of optimal reverse prediction. Semi-supervised learning is of great interest because it can use unlabeled data to improve supervised learning tasks when the labeled data is scarce or expensive.

2.2 Data in intelligent model

The data in modeling algorithms is grouped into online data, offline data, and mixed online and offline data. Reference [18] defines what online and offline data are and introduces the relationship between them.

Most modeling algorithms are based on offline data and models do not change when processes run, which may lead to an instant error due to lack of updating. Two characteristics of offline data for modeling are expected. One is the sufficiency. It is assumed that offline data is the abundance of samples that evenly spread over the range of interests in the sample space. The other is the completeness. The offline data sequence in a sample should be complete or dense almost everywhere.

Many studies on online data have appeared recently, mainly including online passive-aggressive algorithms [19], online kernel learning in a reproducing kernel Hilbert space [20], and implicit online learning with kernels [21]. Online modeling algorithms are more useful when the system to be identified is time-variant, because online data captures the latest information and reflects varying characteristics of a process. However, some drawbacks still exist, for example, it suffers from local information, data loss and bias because online data only contains a portion of a sample.

Based on the above mentioned problems, modeling based on mixed online and offline data is really expected, which can make full use of online data and offline data to overcome some existent difficulties. One idea in Ref. [18] proposed to find the similar relationship between online data and offline data and use offline data to revise online data as though it is online.

2.3 Control methods of data-based model

The main control methods of data-based model include inverse control, internal model control and predictive control, etc. They require that the controlled plant is invertible and stable.

2.3.1 Inverse control of data-based model

Inverse control has been widely approved in the control of nonlinear systems due to evident physical meaning and simple implementation. The inverse model of a controlled system is used as a serial controller in an open-loop way. Instable phenomenon may be aroused due to an open way so that an added controller is introduced to form a closed-loop control with error feedback as shown in Fig. 2.

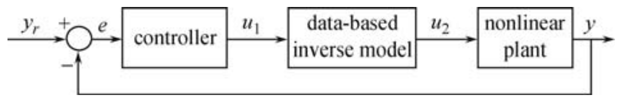


Fig. 2 Schematic diagram of inverse control with error feedback

Consider a reversible nonlinear system described in the following input-output model:

$$F(Y(k+\alpha), Y_\Omega, U(k), U_\Omega) = 0, \tag{1}$$

where $Y \in \mathbb{R}^n$, $U \in \mathbb{R}^m$, $Y(k+\alpha) = (y_1(k+\alpha_1), y_2(k+\alpha_2), \ldots, y_n(k+\alpha_n))$, $Y_\Omega = (y_1(k+\alpha_1-1), \ldots, y_1(k+\alpha_1-p_1), y_2(k+\alpha_2-1), \ldots, y_2(k+\alpha_2-p_2), \ldots, y_n(k+\alpha_n-1), \ldots, y_n(k+\alpha_n-p_n))$, $U(k) = (u_1(k), u_2(k), \ldots, u_m(k))$, $U_\Omega = (u_1(k-1), \ldots, u_1(k-q_1), \ldots, u_m(k-1), \ldots, u_m(k-q_m))$, $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ denote the relative delays of the output to the input, $q = (q_1, q_2, \ldots, q_m)$ and $p = (p_1, p_2, \ldots, p_n)$ express the input orders and the output orders, respectively.

If we set $\phi(k) = Y(k+\alpha)$ as the vector expression of $\phi_1(k) = y_1(k+\alpha_1), \phi_2(k) = y_2(k+\alpha_2), \ldots, \phi_n(k) = y_n(k+\alpha_n)$, the inverse system can be expressed as follows:

$$U(k) = f(Y(k+\alpha), Y_\Omega, U_\Omega) = f(\phi(k), Y_\Omega, U_\Omega). \tag{2}$$

The formula of the input $\phi(k)$ and the output $U(k)$ is the inverse expression of the original system. In the formula, $\phi(k)$, Y_Ω and U_Ω are measurable, so that intelligent modeling algorithms can be applied to approach inverse model based on data. Then the inverse model is cascaded to rebuild a composite linear system with the decoupling transfer function

$$G_{ij}(z) = \frac{y_i(z)}{\phi_j(z)} = \begin{cases} z^{-\alpha_i}, & i = j, \\ 0, & i \neq j. \end{cases} \tag{3}$$

It has a linear transfer function, although nonlinear coupling still exists within the composite system. So the inverse model is considered as a decoupling controller.

The idea of inverse control used in intelligent model can be observed from the following work. Dai et al. presented an NN inverse control method [22], and systematically studied the inverse problem of nonlinear systems based on neural networks. Under the condition of the existent inverse model of a controlled plant, when variables to approach the inverse model can be measurable, neural networks were utilized directly. When unmeasurable variables were needed in approaching, the assumed inherent sensor was proposed to estimate crucial process variables, see Ref. [22], where unmeasurable variables were acquired using the assumed inherent sensor and then the inverse model was approached by neural network. In Ref. [23], an adaptive inverse control algorithm was proposed, where a fast online SVR algorithm was used to approach the inverse model of a controlled plant, then applied the online inverse model as a direct inverse controller, and output errors were used to adjust online SVR inverse model.

2.3.2 Internal model control of data-based model

Internal model control was put forward by Garcia in 1982 with good tracking performance and strong robustness against unknown disturbance and model mismatch [24,25]. Internal model control uses an explicit model in parallel with the process, but most industry processes include nonlinearity, delays and time varying, and they are very difficult to establish explicit models so that data-based identified algorithms are applied to approach controlled plants and their inverse models.

A basic internal model control framework is shown in Fig. 3, where the inversion of internal model is constituted as a controller, and adjusted by error feedback. The internal model and its inversion can be estimated by intelligent algorithms and optimization methods based on data [26]. Internal model control based on data has many research results. Nahas identified the controlled plant and its inversion using BP networks and studied reversible conditions of nonlinear systems, which truly opened up the study of neural network internal model control based on input-output data [27]. Wang et al. used SVR to model the internal model and the inverse controller under the reversible condition, and results showed that the internal model control based on SVR had a simplified model and good control performance compared with that based on neural networks [28]. The 2-port control structure was introduced to internal model control to develop a complete framework in Ref. [29]. The structure can improve the performance of tracking, anti-windup, and anti-interference. Reference [30] presented enhanced internal model control against uncertainties, which consisted of a nonlinear model control and an error feedback loop to achieve disturbance attenuation, but the tracking performance was not modified.

Reference [31] proposed a new internal model control method based on inverse model. This method identified the inverse model and then cascaded the inverse model and the original system to decouple the nonlinear system into several composite linear subsystems. The internal model control was applied to the composite system. The diagram of improved internal model control is shown in Fig. 3.

In the new framework, the internal model is in the form of the linear transfer function, which avoids the situation of irreversible internal model $G_m(z)$. Therefore, the inversion of internal model has analytic expression according to the linear transfer function in Eq. (3), namely $G_m(z) = \mathrm{diag}\{z^{-\alpha_1}, z^{-\alpha_2}, \ldots, z^{-\alpha_n}\}$. The internal model controller $G_c(z)$ is the product of a robust filter $F(z)$ and $G_m^{-1}(z)$ [20]. $d(z)$ is a jamming signal. The tracking errors of the internal model control system can be described as

$$e(z) = y(z) - y_r(z) = \frac{G_c(z)G_m(z) - 1}{1 + G_c(z)(G(z) - G_m(z))} y_r(z).$$

Considering the composite system $G(z)$ may have errors in modeling, the pseudo linear system can be assumed $G(z) = G_m(z)(1+h_m(z))$. The internal model controller can be written as $G_c(z) = F(z)G_m^{-1}(z)$.

2.3.3 Model predictive control of data-based model

Model predictive control is recognized as a very powerful

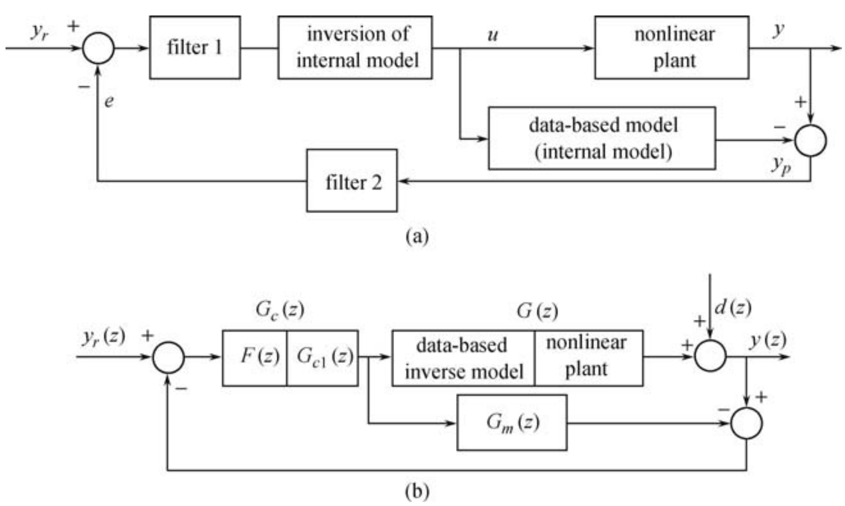


Fig. 3 Diagram of internal model control. (a) A basic framework; (b) an improved framework

approach to handle actual industrial control problems. The main issue in implementation of nonlinear model predictive control is the computational complexity of control variables [32]. The rationale underlying model predictive control is to transform a control problem into an optimization one, so that at any sampling time, a sequence of future control values is computed by solving a finite horizon optimal problem [33–35]. For a nonlinear system, by setting the current values of the system as initial values, the data-based model is used to predict the future outputs over a length of a moving horizon N at each sampling instant, future control values are solved by a finite horizon M optimal problem.

The optimization problem of model predictive control is formulated as

$$\begin{aligned} \min \ J = & \sum_{v=1}^{N}\sum_{i=1}^{n}[y_{ri}(k+v) - y_{bi}(k+v)]^2 \\ & + \sum_{v=0}^{M-1}\eta_v \sum_{i=1}^{m}[u_i(k+v) - u_i(k+v-1)]^2, \\ \text{s.t. } & u_{\min} \leqslant u_i(k+v) \leqslant u_{\max}; \\ & \Delta u_{\min} \leqslant u_i(k+v) - u_i(k+v-1) \leqslant \Delta u_{\max}. \end{aligned} \tag{4}$$

$y_{ri}(k+v) - y_{bi}(k+v)$ expresses errors between predicted outputs and desired trajectories, $u_i(k+v) - u_i(k+v-1)$ expresses the change rate of a control signal, and η_v denotes a weighted term. Only the first element of the computed sequence is effectively used while all subsequent control signals in the control sequence are discarded. This procedure is repeated at every sampling instant using the most updated data of the process [36].

Because of the model mismatch and interference, predictive outputs and actual outputs still have predictive errors. The expression of predictive errors is $e_i(k) = y_i(k) - y_{pi}(k)$. Predictive errors reflect the uncertainty that is not included in the predictive model, which would be a helpful supplement for prediction. Based on the way of feedback compensation to correct future outputs, predictive outputs can be expressed as $y_{bi}(k+v) = y_{pi}(k+v) + he_i(k), v = 1, 2, \ldots, N$. The structure of data-based model predictive control is shown in Fig. 4.

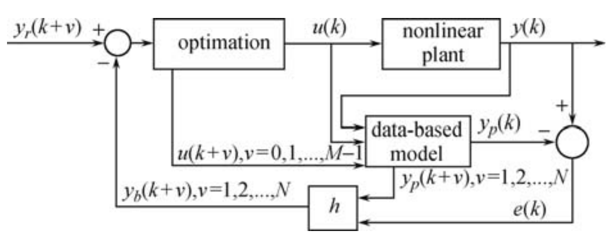


Fig. 4 Structure map of model predictive control

In Ref. [37], continuous time recurrent neural networks were used in nonlinear model predictive control. The future dynamic behaviors of a nonlinear process were predicted by neural networks in real time. Automatic differentiation techniques were developed as an efficient training algorithm. The same neural networks approach was also used to solve the online optimization problem in the predictive controller. In Ref. [38], we presented a single-step predictive control algorithm based on data-based model. The method utilized least square support vector machines to approach a nonlinear system and forecast the outputs and reduced predictive errors by feedback correction. The control values were obtained by the rolling optimization of particle swarm optimization.

3 Data-based controller design

3.1 Outline for direct controller design methods based on data

In this section, we mainly elaborate some existing methods from data to controller directly. Some methods based on online data have simultaneous perturbation stochastic approximation (SPSA), model free adaptive control (MFAC), unfalsified control, and so on. SPSA was presented by Spall in 1993 [39–41] which is a direct approximation control method. This method does not require knowing the model of the controlled plant, and only makes use of closed-loop data to tune parameters of the controller. MFAC was proposed by Hou in 1994 [42,43]. The basic idea is that an equivalent dynamic linear model replaces the general discrete-time nonlinear system at the current operating point, and estimates online the pseudo-partial derivative in the dynamic linear model only with input-output data, so as to realize model free adaptive control. Unfalsified control presented by Safonov is a non-model adaptive control method and selects a specific controller as the current controller from the candidate controllers based on input-output data [44].

Yamamoto proposed data-based proportional-integral-differential (PID) control that tuned PID parameters with data and broadened the traditional PID method [45]. Hjalmarsson proposed iterative feedback tuning that found the optimal controller parameters using the data of closed-loop system and gradient-based iteration. These are based on offline data [46–48].

Control based on offline data cannot make a timely response and adjustment to the adventitious interference such as load variations and other uncertainties. Control based on online data may not meet the condition of sufficient incentive, which leads to a model mismatch and unmodeled errors. The ideal data-based control is always expected to make use of online and offline data during the process on different levels as we mentioned above. Iterative learning control and approximate dynamic programming, as typical algorithms, are proposed based on online and offline data.

3.2 Approximate dynamic programming method for controller design based on data

Dynamic programming is a general approach for sequential multiple stage optimization [49]. Consider a system $x(t+1) = f(x(t), u(t), t)$, where $x \in X \subset \mathbb{R}^n$ is the state and $u \in A \subset \mathbb{R}^m$ is the control action. The performance index is typically a discounted infinite sum of single-stage cost,

$$J(x(i), i) = \sum_{k=i}^{\infty} \gamma^{k-i} U(x(k), u(k), k), \tag{5}$$

where U is the given utility function and γ is the discount factor with $0 < \gamma < 1$. According to Bellman's principle of optimality, the optimal cost from time t is unique and satisfies

$$J^*(x(t), t) = \min_{u(t)} \{U(x(t), u(t), t) + \gamma J^*(x(t+1), t+1)\}. \tag{6}$$

The optimal control $u^*(t)$ at time is the $u(t)$ that achieves this minimum, i.e.,

$$u^*(t) = \arg\min_{u(t)} \{U(x(t), u(t), t) + \gamma J^*(x(t+1), t+1)\}. \tag{7}$$

Approximate dynamic programming has its roots in artificial intelligence and closely follows the ideas of reinforcement learning and neuro-dynamic programming. The approach makes use of online and offline data, estimates the control index or its partial derivatives by an approximate function, and thus optimizes a control law to make it close to the control law given by the traditional dynamic programming. It is characterized by function approximation and iterative improvement of a control law within limited regions of state space. The usual barrier of curse-of-dimensionality [50] is alleviated by employing closed-loop simulations and function approximation. A typical approximate dynamic programming includes model, critic network and action network shown in Fig. 5. Each module can be achieved by neural networks which compose an adaptive dynamic programming system. Critic network and action network are trained by minimizing $E^2(t) = \{Q(t-1) - U(t) - \gamma Q(t)\}^2$ and $U(t) + \gamma Q(t)$.

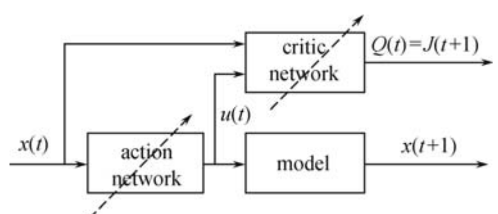


Fig. 5 Two modules in a typical action-dependent approximate dynamic programming

Approximate dynamic programming has also made a lot of success in data-based optimal control. In 2001, Si proposed an online reinforcement learning approach [51]. This method did not require access to the nonlinear mathematic model of the system and utilized neural networks to approximate the control index and optimize the control law according to online data and control signals, which achieved the online optimal control of a nonlinear system. In Ref. [49], the authors presented two different approximate dynamic programming strategies for nonlinear control using input-output data alone. One called J-learning, which built an empirical model and performed dynamic programming with it to derive an improved control policy. The other called Q-learning, which learned an improved control by periodically updating online data without involving an explicit model.

Data-based approximate dynamic programming has penetrated into a large number of scientific fields and has made great progress, but main theories require the known mathematic model of a nonlinear system [52–54] and control systems studied are relatively simple systems. The current data-based approximate dynamic programming only involves the theoretical analysis and simulation and is difficult to directly control actual systems. It still does not consider some difficult factors existing in complex systems, such as nonlinear, time-delay, coupling, and so on.

3.3 Iterative learning control for controller design based on data

Iterative learning control was first presented in Ref. [55] and has been applied in various industrial systems, including robot arm manipulators, chemical batch processes and reliability testing rigs. It shows good performance for the trajectory tracking of unknown plants, which attracts many scholars at home and abroad and have done further study from various angles.

Iterative learning control assumes that the system can repeatedly operate on the same task under the same conditions and the desired output is defined first. The theory mainly includes stability and convergence, learning laws and system structure, learning speed, robustness, analysis methods of iterative learning process, and initial value problems. These elements are all interrelated. Mathematic tools of proving stability and convergence are mainly calculus inequalities, 2-D theory, Lyapunov theory and operator theory. Learning laws contain P-type, D-type, PID-type, higher-order learning law, robust learning law, as well as other learning algorithms. System structure means an open-loop learning, a closed-loop learning, and a combination of open-loop and closed-loop learning. To speed up the learning, an optimal learning control scheme and a higher-order learning law were proposed. The robustness of iterative learning refers to that the iteration can converge to the neighborhood of the desired values under a variety of bounded disturbances and can converge to the desired values when interference is canceled. Robustness is a broader problem than stability [56–59]. Analytical methods of iterative learning mainly include frequency-domain, time-domain, 2-D analysis, and so on.

The initial states in most algorithms are usually set as $x_k(0) = x_d(0)$, $k = 0, 1, 2, \ldots$, and k is the number of iteration. Assuming that a system repetitively performs a given task during a period of time, it can be described as

$$\begin{cases} \dot{x}(t) = f(x(t), u(t), t), \\ y(t) = g(x(t), u(t), t), \end{cases} \tag{8}$$

where $x \in \mathbb{R}^N$, $y \in \mathbb{R}^M$, and $u \in \mathbb{R}^P$ denote states, inputs and outputs, respectively, $f(\cdot)$ and $g(\cdot)$ express functions with unknown parameters and uncertain structure, t denotes different time step and T is the duration. In a given period $t \in [0, T]$, with an initial state $x_k(0)$ and a desired output $y_\mathrm{d}(t)$, the aim is to find a control $u_k(t)$, which makes the outputs $y_k(t)$ track the reference $y_\mathrm{d}(t)$ as close as possible. If we define tracking error as $e_k(t) = y_\mathrm{d}(t) - y_k(t)$, the aim is that $\lim_{k\to\infty} e_k(t) = 0$. The control variables are described as $u_{k+1}(t) = L(u_{k+1}(< t)$, $u_k(\cdot)$, $u_{k-1}(\cdot)$, $\ldots$, $u_{k-n}(\cdot)$, $e_{k+1}(< t)$, $e_k(\cdot)$, $e_{k-1}(\cdot)$, $\ldots$, $e_{k-n}(\cdot))$, which shows that the current control variables depend on previous control variables and previous errors. L is a linear or nonlinear operator. Figure 6 is the block diagram of iterative learning control.

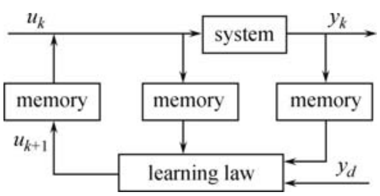


Fig. 6 Block diagram of iterative learning control

Because it utilizes the previous control data and the output error data to modify the current control, iterative learning control is currently only limited to some repetitive processes. However, a large number of industrial systems do not have the character of repeating action, and therefore it is worth exploring to extend iterative learning control to broader industrial systems. Convergence rate is always an important topic for iterative learning control. How to use a priori knowledge and its previous learning information to improve the convergence rate is still a valuable topic.

4 Conclusion and perspective

In this paper, we divide data-based control for nonlinear systems into two subfields, intelligent modeling and direct controller design. We first discuss some algorithms to show the state of the art of data-based modeling, and then some control strategies of data-based model are introduced. Second some approaches to design a controller directly are demonstrated and methods that take advantages of offline data and online data are emphasized. The theory and methods of data-based control for nonlinear systems are only in its infancy; there is still a lot of work to be explored. Some remaining challenges are considered as follows.

1) How to use offline data and online data. Offline data is considered as perfect without interference data and includes all kinds of patterns of the system. The advantage of online data is that they show dynamic characteristics. Several groups of online data within a certain time may express the behavior of a system more accurately. How do we measure the similarities between offline and online data and find similar offline data? It is valuable to weigh the similarities of offline and online data to design a local controller or a local model. The other idea is how to cancel the impact of disturbance data by studying the similarities between offline and online data. If one traverses all offline data and finds that there is no pattern of the online data, the weighted term can be adjusted to reduce the effect of the online data.

2) How to reduce redundant data. Industrial processes produce volumes of data. If all of data is used to model or design a controller, it wastes computational time and memory and seems an impossible work. Therefore, we should find an approach to understand and extract effective data from mass data. How do we justify that the observed data is independent or not and how many categories can approximate the model or a controller without redundancy? From the mathematic point of view, there is no co-linear representation between different dimensions of data and it is a process to find independent vectors.

3) How to determine the stability of data-based control. All information can be considered as data. The theory of data-based control is a very difficult and very powerful problem and still many concepts are not defined well, such as robustness of data-based control. For this new control, many problems are worth exploring. In a closed-loop system of data-based model, if the controlled plant, the data-based model and the controller are stable, is the control system closed-loop stable and how do we decide that the data-based model is stable? How do we determine if the data-based controller is stable and robust?

4) How to acquire the state space model based on data. Most previous methods have been proposed for approaching the input-output model of plants. It is worth thinking how to approach the state space model based on data [60] because of clearer physical meaning, more suitable for representing plants, easier incorporation of

prior physical knowledge into models and smaller repressor number than input-output models.

Although data-based control will come across many challenges, we believe that it has an enviable perspective. Effective control based on mass data is an important direction with the development of information science.

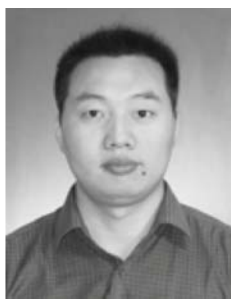
Acknowledgements The authors would like to thank the referees for their valuable suggestions and comments. This work was supported by the National Natural Science Foundation of China (Grant Nos. 60874013, 60953001 and 61034002).

References

1. Brundle M, Naedele M. Security for process control systems: an overview. IEEE Security & Privacy, 2008, 6(6): 24–29
2. Katayama T, Mckelvey T, Sano A, Cassandras C G, Campi M C. Trends in systems and signals: status report prepred by the IFAC coordinating committee on systems and signals. Annual Reviews in Control, 2006, 30(1): 5–17
3. Chai T Y. Challenges of optimal control for plant-wide production processes in terms of control and optimization theories. Acta Automatica Sinica, 2009, 35(6): 641–648 (in Chinese)
4. Lee J H, Lee K S. Iterative learning control applied to batch processes: an overview. Control Engineering Practice, 2007, 15(10): 1306–1318
5. Hou Z S, Xu J X. On data-driven control theory: the state of the art and perspective. Acta Automatica Sinica, 2009, 35(6): 650–667 (in Chinese)
6. Hand D J, Mannila H, Smyth P. Principles of Data Mining. Cambridge: The MIT Press, 2000
7. Han J W, Kamber M. Data Mining: Concepts and Techniques. San Francisco: Morgan Kaufmann, 2006
8. Witten I H, Frank E. Data Mining: Practical Machine Learning Tools and Techniques. 2nd ed. San Francisco: Morgan Kaufmann, 2005
9. Sun C Y, Yu W. Neural network for control, robotics and diagnostics. Neural Computing & Applications, 2008, 17(4): 325–326
10. Hou Z G, Zeng Z G, Sun C Y. Computational intelligence for optimization, modeling and control. Neural Computing & Applications, 2009, 18(5): 407–408
11. Zhu D Q, Shi H. Artificial Neural Network Theory and Application. Beijing: Science Press, 2006 (in Chinese)
12. Deng N Y, Tian Y J. A New Method of Data Mining — Support Vector Machine. Beijing: Science Press, 2006 (in Chinese)
13. Sun C Y, Ju P, Li L F. Fuzzy modeling technique with PSO algorithm for short-term load forecasting. Lecture Notes in Computer Science, 2006, 4223: 933–936
14. Sun C Y, Song J Y, Li L F, Ju P. Implementation of hybrid short-term load forecasting system with analysis of temperature sensitivities. Soft Computing, 2008, 12(7): 633–638
15. Wu Z L, Li C H, Zhu J, Huang J. A semi-supervised SVM for manifold learning. In: Proceedings of the 18th International Conference on Pattern Recognition. 2006, 2: 490–493
16. Adankon M M, Cheriet M, Biem A. Semisupervised least squares support vector machine. IEEE Transactions on Neural Networks, 2009, 20(12): 1858–1870
17. Xu L L, White M, Schuurmans D. Optimal reverse prediction, a unified perspective on supervised, unsupervised and semi-supervised learning. In: Proceedings of the 26th International Conference on Machine Learning. 2009, 1137–1144
18. Xu J X, Hou Z S. Notes on data-driven system approaches. Acta Automatica Sinica, 2009, 35(6): 668–675
19. Crammer K, Dekel O, Keshet J, Shalev-Shwartz S, Singer Y. Online passive-aggressive algorithms. Journal of Machine Learning Research, 2006, 7(8): 551–585
20. Kivinen J, Smola A J, Williamson R C. Online learning with kernels. IEEE Transactions on Signal Processing, 2004, 52(8): 2165–2176
21. Cheng L, Vishwanathan S V N, Schuurmans D, Wang S, Caelli T. Implicit online learning with kernels. Advances in Neural Information Processing Systems, 2007, 19: 249–256
22. Dai X Z, Wang W C, Ding Y H, Sun Z Y. “Assumed inherent sensor” inversion based ANN dynamic soft-sensing method and its application in erythromycin fermentation process. Computers and Chemical Engineering, 2006, 30(8): 1203–1225
23. Wang H, Pi D Y, Sun Y X. Online SVM regression algorithm-based adaptive inverse control. Neurocomputing, 2007, 70(4–6): 952–959
24. Garcia C E, Morari M. Internal model control-1: a unifying review and some new results. Industrial Engineering Chemistry Process Design and Development, 1982, 21(2): 308–323
25. Garcia C E, Morari M. Internal model control-2: design procedure for multivariable systems. Industrial Engineering Chemistry Process Design and Development, 1985, 24(3): 472–484
26. Zhou Y, Chen Q W, Hu W L. New developments of research on internal model control. Control Theory & Applications, 2004, 21(3): 475–482
27. Nahas E P. Nonlinear internal model control strategy for neural network models. Computers Chemical Engineering, 1992, 16(12): 1039–1057
28. Wang D C, Fang T J. Internal model control approach based on support vector machines. Control Theory & Applications, 2004, 21(1): 85–88
29. Stephanopoulos G, Huang H P. The 2-port control system. Chemical Engineering Science, 1986, 41(6): 1611–1630
30. Hu Q, Saha P, Rangaiah G P. Internal model control with feedback compensation for uncertain non-linear systems. International Journal of Control, 2001, 74(14): 1456–1466
31. Song F H, Zheng E H. Nonlinear internal-model control based on support vector machine. Control Theory & Applications, 2008, 25(6): 1067–1071
32. Cueli J R, Bordons C. Iterative nonlinear model predictive control. Stability, robustness and applications. Control Engineering Practice, 2008, 16(9): 1023–1034
33. Tsaia P F, Chub J Z, Janga S S, Shiehc S S. Developing a robust model predictive control architecture through regional knowledge analysis of artificial neural networks. Journal of Process Control, 2002, 13(5): 423–435
34. Huang G S, Dexter L A. Realization of robust nonlinear model predictive control by offline optimisation. Journal of Process Control, 2008, 18(5): 431–438
35. Wang L P, Young P C. An improved structure for model predictive control using non-minimal state space realisation. Journal of Process Control, 2006, 16(4): 355–371

36. Xi X C, Poo A N, Chou S K. Support vector regression model predictive control on a HVAC plant. Control Engineering Practice, 2007, 15(8): 897–908
37. Al Seyab R K, Cao Y. Nonlinear system identification for predictive control using continuous time recurrent neural networks and automatic differentiation. Journal of Process Control, 2008, 18(6): 568–581
38. Mu C X, Zhang R M, Sun C Y. LS-SVM predictive control based on PSO for nonlinear systems. Control Theory & Applications, 2010, 27(2): 164–168
39. Spall J C. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. IEEE Transactions on Automatic Control, 1992, 37(3): 332–341
40. Spall J C, Chin D C. Traffic-responsive signal timing for system-wide traffic control. Transportation Research, Part C: Emerging Technologies, 1997, 5(3–4): 153–163
41. Spall J C, Cristion J A. Model-free control of nonlinear stochastic systems with discrete-time measurements. IEEE Transactions on Automatic Control, 1998, 43(9): 1198–1210
42. Hou Z S. The parameter identication, adaptive control and model free learning adaptive control for nonlinear systems. Dissertation for the Doctoral Degree. Shenyang: Northeastern University, 1994 (in Chinese)
43. Hou Z S. Nonparametric Models and Its Adaptive Control Theory. Beijing: Science Press, 1999 (in Chinese)
44. Safonov M G, Tsao T C. The unfalsified control concept and learning. IEEE Transactions on Automatic Control, 1997, 42(6): 843–847
45. Yamamoto K, Takao G, Yamada T. Design of a data-driven PID control. IEEE Transactions on Control System Technology, 2009, 17(1): 29–39
46. Hjalmarsson H, Gunnarsson S, Gevers M. A convergent iterative restricted complexity control design scheme. In: proceedings of the 33rd IEEE Conference on Decision and Control. 1994, 1735–1740
47. Hjalmarsson H. Iterative feedback tuning: an overview. International Journal of Adaptive Control and Signal Processing, 2002, 16(5): 373–395
48. Hjalmarsson H. Control of nonlinear systems using iterative feedback tuning. In: Proceedings of the American Control Conference. 1998, 2083–2087
49. Lee J M, Lee J H. Approximate dynamic programming-based approaches for input-output data-driven control of nonlinear processes. Automatica, 2005, 41(7): 1281–1288
50. Bellman R E. Dynamic Programming. Princeton: Princeton University Press, 1957
51. Si J, Wang Y. Online learning control by association and reinforcement. IEEE Transactions on Neural Networks, 2001, 12(2): 264–276
52. Abu-Khalaf M, Lewis F L. Nearly optimal control laws for nonlinear systems with saturating actuators using a nerual network HJB approach. Automatica, 2005, 41(5): 779–791
53. Zhang H, Wei Q, Luo Y. A novel infinite-time optimal tracking control scheme for a class of discrete-time nonlinear system based on greedy HDP iteration algorithm. IEEE Transactions on Systems, Man, and Cybernetics — Part B: Cybernetics, 2008, 38(4): 937–942
54. Zhang H, Luo Y, Liu D R. Neural-network-based near-optimal control for a class of discrete-time affine nonlinear systems with control constraint. IEEE Transactions on Neural Networks, 2009, 20(9): 1490–1503
55. Arimoto S, Kawamura S, Miyazaki F. Bettering operation of dynamic systems by learning: a new control theory for servomechanism or mechatronic system. In: Proceedings of the 23rd Conference on Decision and Control. 1984, 1064–1069
56. Chen Y, Wen C, Sun M. A robust high-order P-type iterative learning controller using current-iteration tracking error. International Journal of Control, 1997, 68(2): 331–342
57. Chen Y Q, Wen C Y. Iterative Learning Control: Convergence, Robustness and Applications. Lecture Notes in Control and Information Sciences. Berlin: Springer-Verlag, 1999
58. Amann N, Owens D H, Rogers E. 2D systems theory applied to learning control systems. In: Proceedings of the 33rd Conference on Decision and Control. 1994, 985–986
59. Luca A, Paesano G, Ulivi G. A frequency-domain approach to learning control: implementation for a robot manipulator. IEEE Transations on Industrial Electronics, 1992, 39(1): 1–10
60. Deng H, Xu Z, Li H X. A novel neural internal model control for multi-input multi-output nonlinear discrete-time processes. Journal of Process Control, 2009, 19(8): 1392–1400

Chaoxu MU received her B.S. and M.S. degrees in measuring and control technology from Harbin Institute of Technology and Hohai University, respectively in 2006 and 2009. She is currently a doctoral student majoring in control theory and control engineering in School of Automation at Southeast University. Her research interests include intelligent control and learning algorithms, predictive control, nonlinear control and their applications, and flight control.



Changyin SUN is a professor in School of Automation at Southeast University, China. He received the M.S. and Ph.D degrees in electrical engineering from Southeast University, Nanjing, China, respectively in 2001 and 2004. His research interests include intelligent control, neural networks, SVM, pattern recognition, optimal theory, etc. He has published more than 70 papers. He is an IEEE member and the associate editor of IEEE Transactions on Neural Networks, Neural Processing Letters, International Journal of Swarm Intelligence Research, and Acta Automatica Sinica.